

Weekly Report

Yuxin Ma

08.30.2015 - 09.06.2015

Projects

Spark Workflow System

This week we focused on the implementation of the k -means algorithm. Besides we investigated the monitoring data interface of Spark.

1. Monitoring Data

Another document named *REST API.pdf* describes the available monitoring data that can be extracted from the Spark system. Our next step is to perform visual design on the data.

2. Algorithm

The Python version of the k -means algorithm is now ready to use however it is currently not competitive with the official implementation in Spark Machine Learning Library. We will

Huawei Large Graph Project

The web server Nginx and python WSGI server gunicorn has been deployed on our

Plan for the Next Week

For Spark Workflow System:

- Visual design based on the monitoring data;
- start adding intermediate outputs in the algorithm;
- write the related work.

For Huawei Large Graph Project:

- Assist Fangzhou about the implementation plan.